

TITLE OF THE INVENTION

High Capacitive Density Stacked Decoupling Capacitor Structure

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] Not Applicable.

5

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0002] Not Applicable.

10

BACKGROUND OF THE INVENTION

[0003] The present embodiments relate to semiconductor devices and methods and are more particularly directed to a stacked decoupling capacitor structure.

[0004] Semiconductor devices are prevalent in countless different aspects of contemporary society, and as a result, the marketplace for such devices continues to advance at a fairly rapid pace. This advancement is evident in many respects and relates to semiconductor devices either directly or indirectly as well as the methods for forming such devices. For example, the advancement affects numerous device attributes and increases the need for attention to such attributes during design and manufacturing, where such attributes include device size, reliability, yield, and cost. These aspects as well as others are addressed by the prior art and are further improved upon by the preferred embodiments as detailed below.

15

20

[0005] By way of further background, the preferred embodiments relate to decoupling capacitors formed integrally with an integrated circuit. Decoupling capacitors are used to “decouple” or isolate a signal on a conductor. A common example is connecting a decoupling capacitor between a supply voltage and ground such that any transient or other noise signal affecting the power supply is suppressed by the capacitor, thereby decoupling that signal from other components in the device. While formerly decoupling capacitors were sometimes implemented using separate discrete devices, with the continued advancement of integrated circuits a polysilicon to N-well decoupling capacitor also has been developed. As a capacitor in general, this device includes two conductive plates separated by a dielectric. More specifically and as suggested by its name as a polysilicon to N-well decoupling capacitor, one layer of doped polysilicon forms one of the capacitor’s conductive plates while an N-well, formed within either another well or within a semiconductor substrate, forms the other conductive plate. Lastly, these two plates are separated by a dielectric, which is typically silicon dioxide formed at the same time and from the same material that is also used to form the gate dielectric for various transistors being formed with respect to the same well or substrate.

[0006] While polysilicon to N-well decoupling capacitors have been used successfully in various devices, various drawbacks have been observed in such devices as performance demands increase. Particularly, with increased performance demands, there is a corresponding increase in the need for total decoupling capacitance. Further, the location of decoupling capacitors may be distributed in numerous locations across an integrated circuit die. Thus, when using polysilicon to N-well decoupling capacitors in an attempt to satisfy demands, there is a trade-off in that larger amounts of chip area are required to implement the capacitors, where it is well-known that such area can be quite costly in terms of design and, indeed, in some instances, prohibitive. As another consideration, recall that present polysilicon to N-well decoupling capacitors use the transistor gate dielectric layer as a dielectric between the capacitor plates. However, there also is a trend in the transistor art to decrease the thickness of these gate dielectrics, which correspondingly decreases the amount of capacitive material in the polysilicon to N-well decoupling capacitors. Such a result is undesirable for the capacitor for two reasons. First,

it increases the chance of capacitor leakage because the likelihood of leakage increases as dielectric thickness decreases and the need for decoupling increases. Second, with a thinner capacitive dielectric, there is a greater chance of an electrical short-circuit through the dielectric and, hence, between the capacitor plates. The N-well in such a case may
5 provide some resistance to the current flowing through the short-circuit so as to prevent a total device failure, yet the decoupling effect is lost and, hence, with a sufficient number of such failures then device yield is reduced and/or circuit performance is reduced.

[0007] In view of the above, the present inventors provide below alternative embodiments for improving upon various drawbacks of the prior art.

BRIEF SUMMARY OF THE INVENTION

[0008] In one preferred embodiment, there is a capacitive structure. The capacitive structure comprises a semiconductor base region having an upper surface, a well formed within the semiconductor base region and adjacent the upper surface, a first dielectric layer adjacent at least a portion of the upper surface, and a polysilicon layer adjacent the first dielectric layer. The well, the first dielectric layer, and the first polysilicon layer form a first capacitor and are aligned along a planar dimension. The capacitive structure further comprises a first conductive layer positioned with at least a portion overlying at least a portion of the polysilicon layer, a second dielectric layer adjacent the first conductive layer, and a second conductive layer adjacent the second dielectric layer. The first conductive layer, the second dielectric layer, and the second conductive layer form a second capacitor and are aligned along the planar dimension.

[0009] Other aspects are also disclosed and claimed.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

[0010] Figure 1 illustrates a top view of a stacked decoupling capacitor structure according to the preferred embodiments.

[0011] Figure 2 illustrates a first cross-sectional view of the stacked decoupling capacitor structure of Figure 1.

[0012] Figure 3 illustrates a second cross-sectional view of the stacked decoupling capacitor structure of Figure 1.

[0013] Figure 4 illustrates a schematic of an approximate electrical equivalent of the stacked decoupling capacitor structure of Figures 1 through 3.

DETAILED DESCRIPTION OF THE INVENTION

[0014] Figure 1 illustrates a top view of a decoupling capacitor structure 10 according to the preferred embodiment and built within a semiconductor integrated circuit. Structure 10, as explored further below in connection with Figures 2 and 3, is constructed using numerous layers from which various components are formed. Proceeding generally from a bottom-to-top approach in Figure 1, structure 10 includes a well 12, preferably of an N-type material. Toward the top left and bottom right of structure 10 are two silicide regions 14₁ and 14₂, where active doped regions are underneath regions 14₁ and 14₂, but those active regions are not visible from the perspective of Figure 1. In any event, electrical contact to either of silicide regions 14₁ and 14₂ provides contact, via the respective underlying active doped regions, to well 12. A doped polysilicon region 16 is visible toward the left of Figure 1, and for reasons discussed later will be shown to extend a considerable distance in what is shown as the horizontal dimension in Figure 1, although the span of that extension is not visible from the perspective of Figure 1. Four contacts 18₁, 18₂, 18₃, and 18₄, are shown using dotted squares in Figure 1. The dotted convention is used with contacts 18₁, 18₂, 18₃, and 18₄ because such contacts are not actually viewable from the perspective of Figure 1 because they are below the bottom layer of a metal-interlevel oxide-metal ("MIM") capacitor 20, where only a portion of the top layer of MIM capacitor 20 is visible in Figure 1. Structure 10 also includes four metal-1 regions 22₁, 22₂, 22₃, and 22₄, where contacts extend below the first three of those four regions. Metal-1 region 22₁ includes four contacts 24₁, 24₂, 24₃, and 24₄, metal-1 region 22₂ includes four contacts 24₅, 24₆, 24₇, and 24₈, and metal-1 region 22₃ includes four contacts 24₉, 24₁₀, 24₁₁, and 24₁₂. Lastly, structure 10 includes a metal-2 region 26, where eight contacts extend below that region. Specifically, these contacts include contacts 28₁, 28₂, 28₃, 28₄, 28₅, 28₆, 28₇, and 28₈.

[0015] Figure 2 illustrates a cross-sectional view of structure 10 along line 2-2 in Figure 1. Thus, Figure 2 illustrates various of the components discussed above with respect to Figure 1, as well as additional aspects and components that are appreciated from a cross-sectional perspective. Well 12 is formed in a fixed position relative to some

type of semiconductor region, where that region is shown as a region 30 in Figure 2. Such a region could be, by ways of example, another well or a semiconductor substrate and, thus, for sake of reference is hereafter identified as substrate 30. Thus, the remaining illustrated devices also are formed in a fixed position relative to substrate 30. Note also
5 with respect to Figure 2, as well as in Figure 3 described below, that a few simplifications are made as will be appreciated by one skilled in the art so as to not unduly complicate the figures, while still presenting an understanding of the preferred embodiments. First, the figures are not drawn to scale. Second, certain intermediate dielectric or insulating layers are not shown but will be understood to be included so as to provide adequate isolation as
10 well as physical support for the various structures as they are formed in a stacked fashion to create the overall structure 10.

[0016] Returning to well 12, within it is an isolation region 32, sometimes referred to as a shallow trench isolation ("STI") region, which is an insulating material such as an oxide. Also within well 12 is an active region 34₁, and since the present example for well
15 12 is as an N-type material, then preferably active region 34₁ is an N⁺ region, that is, it has a higher concentration of N-type dopants as compared to well 12. Overlying active region 34₁ is silicide region 14₂, formed by applying and reacting a metal and thereby facilitating a lower sheet-resistance electrical contact to active region 34₁. Overlying a portion of the upper surface of well 12 is a gate insulator 38, formed of an insulating substance such as
20 silicon dioxide. The term "gate" is used in connection with this insulator because, in the preferred embodiment, gate insulator 38 is formed at the same time as are other gate insulators for transistors (not expressly shown) that are formed in the same step as the formation of gate insulator 38. In other words, in the preferred embodiment, it is contemplated that structure 10 is formed as part of an integrated circuit that will include a
25 number of transistors, although such transistors are not shown so as to simplify the figures and to focus on the inventive capacitive structure. At the same time as certain steps are used to create those transistors, the same processing steps can be used to form various aspects of structure 10. For example, such structures include the already-described well 12, isolation region 32, active region 34₁, silicide region 14₂, and gate insulator 38. Indeed,
30 one skilled in the art will appreciate that many of the later-described-regions also may be

formed at the same time and using the same steps as are known and used in the art to form integrated circuit transistors and their interconnects. In any event, therefore, when the gate insulators are formed for such transistors, at the same time gate insulator 38 is formed.

5 [0017] Continuing with structure 10 as shown in Figure 2, overlying gate insulator 38 and isolation region 32 is doped polysilicon region 16. Recall in Figure 1 that only a portion of doped polysilicon region 16 is visible toward the left of that Figure. However, in Figure 2, it may be seen that doped polysilicon region 16 in the preferred embodiment extends a majority of the horizontal distance across structure 10. The function of this extension is now better understood in that the combination of doped polysilicon region 16, gate insulator 38, and well 12 form a first capacitor in structure 12. More particularly, doped polysilicon region 16 serves as a first capacitor plate, well 12 serves a second capacitor plate, and gate insulator 38 between them serves as a capacitor dielectric. Thus, it may be seen that each of these three layers generally extends in a parallel planar dimension and that plane is also parallel to a plane through substrate 30. Given this first capacitor, the greater the area covered by doped polysilicon region 16 and the corresponding dielectric covered by gate insulator 38, the larger the device capacitance. Such an increase is highly desirable for a decoupling capacitor, as is the purpose served by structure 10.

20 [0018] Continuing still further with structure 10 as shown in Figure 2, contact 18₃ is shown between doped polysilicon region 16 and MIM capacitor 20, where MIM capacitor 20 provides a second capacitor in structure 10. Specifically, from the perspective of Figure 2, it may be seen that MIM capacitor 20 includes three layers 20₁, 20₂, and 20₃, and each of these layers also is generally along and parallel to the same planar dimension along which substrate 30 and the above-described capacitor, formed by polysilicon layer 16, gate insulator 38, and well 12, are aligned. Further, contact 18₃ contacts the lower of those layers, namely, layer 20₁, where therefore contact 18₃ is substantially perpendicular to the plane spanned by the two above-discussed capacitors. Layers 20₁ and 20₃ are conductive plates, while layer 20₂ is a dielectric, all of which are detailed later. At this point, however,

note then that an electrical bias, as applied to doped polysilicon region 16, is communicated by contact 18₃ to layer 20₁, that is, the electrical bias is connected to one plate of the second capacitor provided by MIM capacitor 20. Overlying layer 20₃ of MIM capacitor 20, that is, the upper layer of that capacitor, is metal-1 region 22₄. At the same time metal-1 region 22₄ is formed, so are metal-1 regions 22₁ and 22₃. Metal-1 region 22₁ is electrically connected by contact 24₂ to doped polysilicon region 16. Note two additional observations with respect to this connection. First, preferably any contact to doped polysilicon region 16 is preferably made over an isolation region, and it is for this reason that isolation region 32 was formed in well 12, that is, due to the subsequent need to contact doped polysilicon region 16 with contacts 18₃ and 24₂. Second, given the illustration as has now been described, note further by way of example that a voltage of 3 volts is shown connected to metal-1 region 22₁. Thus, the 3 volts connected to metal-1 region 22₁ is electrically communicated by contact 24₂ to doped polysilicon region 16, and from doped polysilicon region 16 that 3 volts is electrically communicated by contact 18₃ to layer 20₁, that is, to the to the bottom plate of MIM capacitor 20. Lastly, metal-1 region 22₃ is electrically connected by contact 24₁₂ to silicide region 14₂ which, as described earlier, is electrically connected to active region 34₁. Finally, for reasons detailed below, note that Figure 2 illustrates that metal-1 region 22₃ is connected to ground. Thus, this ground potential is electrically communicated by contact 24₁₂, silicide region 14₂, and active region 34₁, to the N-type well 12.

[0019] Figure 3 illustrates a cross-sectional view of structure 10 along line 3-3 in Figure 1 and, again, therefore, it illustrates various of the components discussed above with respect to Figure 1. Figure 3 also illustrates additional aspects and components that are appreciated from a cross-sectional perspective and that are not shown in the cross-sectional perspective of Figure 2. However, to the right of Figure 3, it illustrates generally the same components shown in Figure 2 and, thus, the reader is either assumed familiar with the earlier discussion or is referred to that discussion for additional detail. Accordingly, the following discussion focuses on the items shown in Figure 3 that are not shown in Figure 2.

[0020] Looking to the left of Figure 3 and starting at the top of metal-1 region 22₄, it is electrically connected by two contacts 28₇ and 28₈ to metal-2 region 26. Further, metal-2 region 26 is electrically connected to metal-1 region 22₂ by two contacts 28₃ and 28₄. Still further, metal-1 region 22₂ is electrically connected to silicide region 14₁ by a contact 24₇, and silicide region 14₁ overlies an active region 34₂, which in the preferred embodiment is formed at the same time as active region 34₁ of Figures 2 and 3 and, thus, which is an N⁺ region having a higher concentration of N-type dopants as compared to well 12. Having described these various connections to the left of Figure 3 and for introducing an aspect described later, note now the overall electrical path created by them. Specifically, recall that layer 20₃ represents a top capacitor plate for MIM capacitor 20. Further, layer 20₃ is in physical and electrical contact with metal-1 region 22₄, which through contacts 28₇ and 28₈, is in electrical contact with metal-2 region 26, which through contacts 28₃ and 28₄, is in electrical contact with metal-1 region 22₂, which through contact 24₇, silicide region 14₁, and active region 34₂, is in electrical contact with N-type well 12.

[0021] Having described structure 10 from various perspectives, attention is now directed to a preferred manner in which it may be constructed. By way of introduction, recall earlier in connection with gate insulator 38 that the discussion states that various of the components of structure 10 are preferably formed using the same steps that are used to form transistors, also associated with substrate 30. As such and as now further appreciated, structure 10 can be implemented with very little alteration to an existing transistor process flow, thereby improving overall device cost. Generally, therefore, up to the formation of MIM capacitor 20, all devices shown below the horizontal plane of MIM capacitor 20 in Figures 2 and 3 are constructed according to principles used for the formation of prior art transistors. For example, active regions 34₁ and 34₂ are formed at the same time as are the source/drain implants for any nearby transistor(s) and those areas are silicided at the same time as the transistor source/drains, gate insulator 38 is formed at the same time as is the gate insulator for any nearby transistor(s), and doped polysilicon region 16 is formed at the same time as is the gate for any nearby transistor(s). Thereafter, an insulating layer (not shown) is formed, patterned and etched to form holes through

that layer, and those holes are filled with a conductive material, such as tungsten, to form contacts 24₇ and 24₁₂.

[0022] In the preferred embodiment, after contacts 24₇ and 24₁₂ are formed, MIM capacitor 20 is formed, where its formation entails steps not necessarily provided by the construction of transistors or the interconnectivity of those transistors. In the preferred embodiment, MIM capacitor 20 is formed by first forming a first conductive layer, which in the preferred embodiment is achieved by depositing a conductive layer, such as tantalum nitride (TaN) in the preferred embodiment. Next, a dielectric layer is deposited over the first conductive layer, such as Ta₂O₅ in the preferred embodiment. Note that Ta₂O₅ has heretofore been used in the dynamic random access memory ("DRAM") art and therefore in the preferred embodiment the process used in that art may be employed here, albeit to create an inventive overall structure 10. Further, in alternative preferred embodiments, a material other than Ta₂O₅ may be used for the dielectric layer, where preferably such a material has a sufficiently high dielectric constant so as to provide a greater decoupling effect, where a desirable dielectric constant may be greater than 4.0. Next, a second conductive layer is deposited over the dielectric layer, again such as TaN in the preferred embodiment. At this point, the three layers are patterned and etched at the same time, thereby forming a same dimension uniform stack, providing MIM capacitor 20.

[0023] Following the formation of MIM capacitor 20, again the remaining portions of structure 10 can be implemented according to principles used for the formation of prior art transistors and the interconnectivity between them, thus requiring no alteration to existing process flow other than, of course, having formed MIM capacitor 20 prior to the formation of metal-1 as described above. For example with respect to the remaining process, following the formation of MIM capacitor 20, an intermetal dielectric can be formed (not shown) and thereafter all metal-1 regions can be formed, that is, the first level of metal proximate substrate 30 is formed so as to provide local and possibly global interconnects. This step will include the formation of metal 1 regions 22₁ through 22₄. In the preferred embodiment, however, note that metal-1 region 22₄ preferably has surface dimensions that are less than those of MIM capacitor 20. In other words, looking to Figure 1, it may be

seen that MIM capacitor 20 has a surface length, L , in a first dimension (shown horizontally) and a surface width, W , in a perpendicular dimension (shown vertically). When metal-1 region 22₄ is formed, its corresponding length and width are preferably less than L and W , respectively. Returning to Figure 2, the reason for this preference is to prevent metal-1 region 22₄ from extending beyond an edge 20_E of MIM capacitor 20, for if such an extension were to occur, it may short-circuit layers 20₁ and 20₃ together, thereby bypassing the capacitive effect provided by MIM capacitor 20. In any event, once the metal-1 layer is complete, the remaining layers and connections, that are still farther from substrate 30 than the metal-1 layer regions, also may be achieved as known in the transistor art. For example, a dual damascene process may be used for the formation of a metal-2 layer, thereby forming metal-2 region 26 and contacts 28₁ through 28₈ at the same time and from the same material, such as copper, or alternatively, contacts 28₁ through 28₈ may be formed first, followed by the formation of a metal-2 layer which can be patterned so as to provide metal-2 region 26. Also in this regard, note that contacts 28₁ through 28₈ are sometimes referred to as vias.

[0024] Figure 4 illustrates a schematic of an approximation of the circuitry realized by structure 10, as will now be more apparent given the previous illustrations and discussion of the various components. In general, Figure 4 demonstrates that there are two parallel capacitors. Looking to the first capacitor of Figure 4, shown to the left in that figure, and with reference to Figure 2, that first capacitor has one plate, namely doped polysilicon region 16, connected to 3 volts, as is achieved in an electrical path that passes through contact 24₂ and metal-1 region 22₁. The second plate of the first capacitor is provided by the N-type well 12; further, however, note from Figures 1 through 3 that well 12 extends a considerable horizontal distance, where that distance in the preferred embodiment is at least 25% of the dimension L spanned by the surface length of MIM capacitor 20. Thus, the distance spanned by N-type well 12 also provides a resistance, and the right end of that resistance is connected to ground, as shown in the schematic of Figure 4 and also as achieved in Figures 2 and 3 through an electrical path that passes through active region 34₁, silicide region 14₂, contact 24₁₂, and metal-1 region 22₃. Looking now to the second capacitor of Figure 4, shown to the right in that figure, and with reference to Figure 2, that

second capacitor is MIM capacitor 20. Further, MIM capacitor 20 has a first plate, namely layer 20₁, that is connected to 3 volts through an electrical path that passes through contact 18₃, doped polysilicon layer 16, contact 24₂, and metal-1 region 22₁. In other words, the positive voltage is applied to MIM capacitor 20 from its underside, that is, from the direction toward substrate 30. Additionally, MIM capacitor 20 has a second plate, namely layer 20₃, separated from layer 20₁ by a dielectric layer 20₂, and that layer 20₃ is connected through a resistor to ground. Specifically with reference to Figure 3, layer 20₃ is connected in an electrical path that passes through metal-1 region 22₄, contacts 28₇ and 28₈, metal-2 region 26, contacts 28₃ and 28₄, metal-1 region 22₂, contact 24₇, silicide region 14₁, and active region 34₂ to the left end of N-type well 12. Again, the horizontal span of well 12 between its left and right ends provides a resistance between those ends that is connected to ground at its right end, through an electrical path that passes through active region 34₁, silicide region 14₂, and metal-1 region 22₃.

[0025] From the above, one skilled in the art should appreciate that the preferred embodiments provide for a high capacitive density stacked decoupling capacitor structure that includes two capacitors in a vertically stacked configuration. In general, this stacked structure can be seen in Figures 2 and 3 such that each capacitor plate has a length along a same parallel plane (into the page) and such that the plates of the polysilicon-to-N-well capacitor are parallel to substrate 30 and the plates of MIM capacitor 20 are above and parallel to the plates of polysilicon-to-N-well capacitor. Further, from Figure 4, one skilled in the art will appreciate that the two capacitors of structure 10 are essentially in parallel in an electrical sense, in which case the capacitance provided by each capacitor is added for a total overall capacitance. Note that the parallel connection is characterized as essentially parallel and the schematic of Figure 4 is described above as an approximation because the actual resistance seen by each of the two capacitors will differ. Specifically, MIM capacitor 20 is connected to the full length of the N-type well 12 and thereby is connected to that full resistance, while the polysilicon-to-N-well capacitor formed below MIM capacitor 20 uses the N-type well 12 both as a capacitor plate and as a resistance to ground, thereby experiencing less resistance than provided to MIM capacitor 20. In any event, the electrical connection of the two capacitors may be approximated as a parallel connection.

Thus, the first capacitor, which by itself is known in the art, may provide a capacitive density on the order of $5 \text{ fF}/\mu\text{m}^2$, yet that capacitance is added to the capacitive density of the MIM capacitor 20, which preferably is greater such as on the order of $10 \text{ fF}/\mu\text{m}^2$, thereby providing a total capacitive density on the order of $15 \text{ fF}/\mu\text{m}^2$. Moreover, this combined capacitance is achieved using the same amount of surface area as compared to that which would be used if only the polysilicon-to-n-well capacitor were implemented, that is, by stacking the two capacitors in a vertical fashion and thereby avoiding the consumption of additional lateral surface area. In addition, note that the resistance provided by the N-type well 12 serves a protective function for both capacitors. Specifically, should there be a failure of a capacitor dielectric (i.e., either gate insulator 38 or layer 20_2), then there is an electrical short between the two associated capacitor plates. However, rather than that short-circuit connecting current directly to ground, the current is dissipated through the length of the resistance of well 12. As such, in the case of only a single or perhaps a few such failures on a single integrated circuit, the overall circuit still may be usable, albeit, without the decoupling capacitance that otherwise would be provided by the failed capacitor(s). To the contrary, were this resistive protection not provided, then a short-circuit directly to ground might render the entire integrated circuit unusable, thereby substantially reducing device yield in the case of mass fabrication of devices using such an approach. In addition, structure 10 provides these benefits with the additional benefit that it may be constructed with little alteration to the process flow used to construct transistors on the same integrated circuit. Indeed, as described above, only a single additional mask is required, that is, to pattern and etch the stack of layers 20_1 , 20_2 , and 20_3 , once those layers have been formed. However, note that rather than contacting layer 20_1 from its underside (i.e., from the direction of substrate 30), in an alternative embodiment layer 20_1 could be made to extend further laterally than layers 20_2 and 20_3 , in which case layer 20_1 could be contacted from the direction of the metal-2 layer, that is, from what is the top of structure 10 as shown in Figures 2 and 3; however, in such a case, two masks would be required, a first to etch layer 20_1 to extend in this fashion, and a second to etch layers 20_2 and 20_3 back so that a portion of layer 20_1 is exposed from the upper surface of structure 10. Nonetheless, this alternative also provides at least the

benefits of greater capacitive density as well as the resistive fail-safe mechanism of the N-type well 12 in the event of a short circuit for either capacitor. Further, while certain preferred materials have been described, one skilled in the art may ascertain various alternatives that also may be implemented within the inventive teachings. Additionally,

5 while the preceding embodiment has been shown as formed with a certain specified manufacturing process, various steps in that process may be altered and still fall within the inventive scope. Thus, the preceding benefits as well as the various alternative steps described and ascertainable by one skilled in the art demonstrate the flexibility of the inventive scope, and they should also demonstrate that while the present embodiments

10 have been described in detail, various substitutions, modifications or alterations could be made to the descriptions set forth above without departing from the inventive scope which is defined by the following claims.